An Elementary Mathematical Model for the Interpretation of Precipitation Probability Forecasts

J. H. CURTISS

University of Miami, Coral Gables, Fla.

(Manuscript received 5 September 1967, in revised form 30 October 1967)

ABSTRACT

A mathematical model is constructed for interpreting precipitation probability forecasts which have areal connotations. The paper is chiefly concerned with a simple discrete form of the model, in which a forecast area is represented by N rain gages scattered throughout it, and a precipitation event is identified with the observing of more than a trace of precipitation in a particular subset of the N gages. The basic parameters in the model are: (a) conditional probabilities of the events consisting of the selection of particular subsets of the N gages by precipitation, given that a certain proportion r of the gages are selected; and (b) a probability distribution for r. The probabilities in (a) are assumed to be peculiar to the particular forecast area and can easily be estimated from historical data. The probabilities in (b) (or at least the mean value of the distribution) are intended to be the objects of estimation in each forecast. It is apparently customary for forecasters to arrive at a forecast precipitation probability by calculating the theoretical mean value of r and assuming that this is equal to a point probability valid uniformly at each point of the forecast area. As explained within the model, this forecast method actually gives the arithmetic mean of the "point probabilities" of the events that the Jth gage receives more than a trace for $J=1, 2, \cdots, N$. Questions concerning biases in published verification data and possible lack of randomness in sequences of precipitation verifications within probability categories are raised as suggestions for further study concerning the operational significance of precipitation probabilities.

1. Introduction

The idea of expressing a precipitation forecast in terms of numerical probabilities has been under discussion for many years (Cook, 1906; Hallenbeck, 1920). However, it is only rather recently that such probabilities have been included routinely in U. S. Weather Bureau local forecasts. A local forecast pertains to a geographical region which typically includes all points lying in a circle of radius at least 25 mi about the center of an urban area. Thus, a local forecast area generally covers more than 1900 mi^2. A probability forecast appears in the news media in a format such as, "Shower probability, 40 per cent." Usually a single number is given for the entire region. But sometimes an areal or temporal qualification is published along with the probability, as is exemplified by the following forecast quoted verbatim from the *Miami* (Florida) *Herald* for Sunday, 4 June 1967:

Today's forecast

> Miami and vicinity: Partly cloudy through Monday with scattered showers most likely during night and morning hours except well inland during afternoons. High today in eighties. Variable mostly easterly winds 5 to 15 mph. Shower probability 50 per cent.

Here the "50 per cent" seems to be qualified both as to time and as to location.

It is natural for a scientist or mathematician who is familiar with quantitative probability theory, but who is not a meteorologist, to wonder whether such a numerical probability statement can be given a meaning within this theory. The genesis of the present paper was pure curiosity on the part of the author, who has no formal training in meteorology.

The fact is that in the absence of authoritative information as to the intent of the U. S. Weather Bureau, and under the assumption that meaningful quantitative *a priori* probabilities really can be assigned to precipitation events, there are quite a number of reasonable conjectures which can be made as to what a single-number precipitation probability might mean. For example, it might mean: (a) the probability of the event that some rain will fall somewhere in the forecast area sometime during the time period covered by the forecast; (b) the probability of the event that general rain will cover all of the area; (c) the fraction of the forecast area which will receive rain in the forecast period; (d) the probability that a specific point in the forecast area will receive more than a trace of rain sometime during the forecast period.

Preliminary inquiries directed by the author to colleagues knowledgeable in probability and statistical theory yielded a diversity of opinions as to what the U. S. Weather Bureau precipitation probabilities mean, but "nothing" was all too frequently the answer. Surprisingly, in a small sample of atmospheric scientists and meteorologists, there was no consensus. For example, a distinguished weather research scientist told the author he thought that it was the probability of the event (a). However, it is soon clear to anyone who dips into the weather forecasting literature that the official Weather Bureau interpretation is the probability referred to in (d).

Actually the difference between the probabilities (a) and (d) can be very great under certain circumstances.[1] For example, suppose that a forecast is based on the expectation of a particular circulation regime, which, when it has occured in the past, has on the average deposited precipitation on 10% of the forecast area in a 12-hr period, and every time the regime occurred, some precipitation fell somewhere. Suppose further that the probability that this regime will obtain during the forecast period is estimated to be 90%, and the alternative is no rain at all. The analysis in Section 5 shows that if the probability (d) is assumed to be uniform over the forecast area, then this probability in the present case is 9%, whereas the probability (a) is 90%.

Stated in detail, here is the official viewpoint of the Weather Bureau toward precipitation probability forecasts.[2] The event to which a precipitation probability applies is the occurrence of more than a trace of precipitation (water equivalent, if frozen) within a specific forecast period at a specific point in the forecast area. ("More than a trace" at a point means that 0.01 inch or more of precipitation is deposited in a rain gage at the point during the forecast period.) For purposes of verification the specific point is taken to be the location of the rain gage at the official local Weather Bureau Station; for example, this is Midway Airport for Chicago and vicinity. When a single unqualified probability number is released in a local forecast, the assumption is made implicity that local conditions will impose only one regime of events in the metropolitan area, and that the published "point probability" is therefore at least approximately valid at each point of the forecast area.

Thus, a forecaster would be meeting government specifications, so to speak, if he concentrated on estimating the relative frequency, given the forecast information, of the event that more than a trace of rain will be recorded merely in the one gage at the official station. Now if a given forecast is based on the assumption that *either* a general rain will cover the forecast area completely during the forecast period *or* there will be no rain at all, then there can be no ambiguity in the meaning of the forecast probability *for a user who is advised of the assumption*. The probability number released will be the estimated probability of the general rain, and this will obviously be the same as the user's own point probability, wherever he may be located in the forecast area. But suppose that the alternatives to "no rain" include various possibilities as to *partial* coverage of the forecast area by precipitation during the forecast period. Then in order to interpret a single-number precipitation probability forecast as a point probability uniformly valid over an area of at least 1900 mi^2, clearly some assumptions must be made as to the way in which the *location* of any partial coverage will be distributed in the forecast area.

As far as the author is aware, these assumptions as to the probability distribution of covered area have not been discussed explicitly and in generality in the literature of precipitation probability forecasting. It is the purpose of this paper to set up an elementary mathematical model within which the necessary assumptions and various other logical problems attendant upon precipitation probability forecasting for an area of substantial content can be studied. One aspect of the model may be found to be useful by weather scientists in that it provides a procedure for the routine preparation of more than a single point probability in a given forecast, on the basis of a forecast only of the probability distribution of the proportion of area which will be covered by precipitation.

Most of the discussion in the text of the paper is concerned with a simplified discrete form of the model. Here the only observables are represented by a set of dichotomous variables, each of which is given the value 1 if the particular rain gage to which it is assigned shows more than a trace, and 0 otherwise. The exposition uses only elementary mathematical terms which are widely accepted within the theory of statistics. There are no *a priori* assumptions requiring the use of special probability distributions, such as the uniform distribution or the Poisson distribution.

One last word of introduction. The developments in this paper are based on the assumption that standard probability theory really can be applied usefully to precipitation forecasting. It is a favorite practice of mathematicians to side-step really difficult problems by making formal hypotheses. In the author's opinion, there may be such a problem in probability forecasting. It is a problem which is hard even to formulate exactly. An attempt can be made with this question: Are the precipitation "probabilities" which are being released to the public really interpretable as action probabilities in day-to-day decision-making, or are they only some sort of very long-term average frequencies which are inapplicable to a particular single forecast period? The question

[1] In general, the authors of reports and published papers on precipitation probabilities in the meteorological literature are clearly aware of this fact. For example, see Hughes (1965). As another example, Epstein (1966a) sets up an explicit but rather artificial model involving coverage by "random" showers of circular shape and of constant radius. Within this model he derives the exact relationship between the probabilities of (a) and (d).

[2] See, for example, Hughes (1965). The wording used here in the text follows very closely a formulation in a letter to the author dated 13 March 1967, from R. H. Simpson, Associate Director of the U. S. Weather Bureau.

will be discussed in more detail in Section 6, but no real effort will be made to come to grips with it in this paper.[3]

2. A sample space for the precipitation "experiment"

First, a summary of the premises scattered through the introduction will be given.

The type of precipitation forecast under discussion is a local forecast applicable to a forecast area of substantial areal content, typically more than 1900 mi^2. (The discussions will also apply by trivial specialization to a forecast area which is only a pin-pointed spot such as an air field or the local official weather station, but it is doubtful if this paper carries much useful information for such a situation.) "Forecast area" will henceforth be abbreviated to FA. The precipitation forecast applies to a particular period of h hours in the more or less immediate future which will be called the forecast period. Typically, $h=12$ hr.

It is also supposed that the event which a precipitation forecast refers to in general is the wetting of a certain geographical subregion of FA by rain at some time in the forecast period. The rain may fall in one shower, or several intermittant showers, or throughout the forecast period. The amount of rain at a given location will not be taken into consideration, provided that it is not less than 0.01 inch as measured by a rain gage. This threshhold amount will be referred to frequently as "more than a trace."

For the convenience of the reader who does not have a mathematical-statistical backgound, a brief and very incomplete description of the standard modern approach to elementary probability model-building will be sketched in the next three paragraphs. Various detailed treatments are available in the newer probability text books; one which has become a classic is that of Feller (1957).

The starting point is the idea of a *random experiment*, which in the purely mathematical sense is an undefined notion. In general terms, a random experiment consists of some actions performed on a given system, or reactions which automatically take place within a given system, which yield an observable result, but the result is not categorically predictable in advance of the experiment. It is customary to call a result of a random experiment an *event*. Thus, if the random experiment consists of the toss of five coins, more than three falling with heads up is one of the various events. If the experiment consists of dealing the cards in a bridge game, the appearance of a given set of 13 cards in North's hand is one of the various events.

The assignment of probabilities to a random experiment consists in assigning numbers lying in the interval [0, 1] to each of the possible events. The number given to any event is supposed to measure in some way the degree of certainty of this event. The rationale underlying the specific numerical assignment of the probability measure for any given random experiment belongs partly to the philosophical part of probability theory and partly to the realm of statistical theory. There are also some mathematical technicalities which limit the types of events to which probabilities can be assigned.

The concept of the *sample space* for a random experiment is basic to an assignment of probabilities. The sample space is the set of all so-called *simple events*, by which is meant the set of all reasonably possible, different, non-decomposable results of the experiment. An *event* is then a set of simple events, which may consist of only a single simple event. An event is said to *occur* in a performance of the experiment if the outcome of the experiment can be identified as one of the simple events in the event. It is implied that no two simple events can occur simultaneously; simple events are "mutually exclusive." (In the older literature, the simple events comprising a given event are called the "ways in which the given event can occur.") For example, in a bridge deal, an appropriate sample space consists of the set of all different divisions of the bridge deck of 52 cards into four equal piles of 13 cards at four distinct locations, say E, S, W, N. Each such deal is a simple event and is thought of as an element or point in the sample space. The event that N gets all the spades is comprised of all the simple events in which E, S, W are dealt piles containing no spades. For a toss of two coins, a penny and a dime, the sample space might conceivably be represented by the set of symbols {HH, HT, TH, TT, EH, ET, HE, TE, EE}. Here H or T standing first in a symbol means, respectively, that head or tail showed on the penny and E means it stood on edge. There is a similar significance for the dime in the second letters. But, of course, this sample space would usually be simplified by leaving out the symbols with an E in them.

In the case of precipitation, the random experiment consists of the interaction of those meteorological phenomena which affect precipitation over FA during the forecast period. In view of the premises stated at the beginning of this section, the result of a performance of the experiment is merely that each point of a certain subset of the geographical points in FA received more than a trace of precipitation. Thus, the set of all simple events (that is, the sample space) can be identified with the collection $\mathcal{P}$(FA) of all subsets of points in FA.[4] The particular event E in $\mathcal{P}$(FA) occurs if all the points in E and no other points in FA are observed to receive more than a trace in the forecast period. For purposes of

[3] The question of the credibility of probability forecasts has been studied by Epstein (1966b), but he starts with the assumption that the forecasted probability is indeed a probability in some recognized sense (in his case, a subjective probability) which may require modification via Bayes' Law by the user in the light of his experience. This is not the same as the problem posed in the text above, which pertains to the *statistical regularity* of a sequence of observations at the official rain gage, given a particular probability forecast such as 30%.

[4] Mathematicians call $\mathcal{P}$(FA) the power set of the point set FA.

identification as simple events, two subsets E_1 and E_2 of FA, itself considered as a geographic point set, are regarded as "different" if they are unequal in the point set theory sense; that is, if E_1 contains one or more points not contained in E_2, and/or vice versa. Thus, E_1 and E_2 may overlap extensively in the geographical sense but still be different and therefore "mutually exclusive" as to occurrence.

One might be tempted to inquire here as to why such an abstract sample space is required. Would it not suffice to consider just the collection of all geographical points in FA as constituting the sample space? The understanding would be that a point (simple event) occurs if it is observed to receive more than a trace in the forecast period. But a moment's thought reveals that such a sample space would be totally irrelevant to the precipitation random experiment, because the basic idea in the physical interpretation of a sample space is that the simple events in it occur in a mutually exclusive manner. If one of the points in FA were to occur, theoretically none of the other points in the FA could occur during the same forecast period if this sample space were to be adopted.

Theoretically, a probability distribution can be associated with an abstract sample space like $\mathcal{P}$(FA) [see Kolmogorov (1950, p. 46)]. But obviously when doing so in practice, one should somehow take "sizes" of the respective wetted subsets into account. This suggests that it would be useful to stratify the sample space by grouping the subsets into classes according to the amount of FA contained in each subset. This leads to a revised notation for the sample space, i.e., the sample space is represented by the set of all ordered pairs (r, E_r), where r is any number in the interval [0.1] and represents the proportion of FA receiving rain, and E_r is any point set which contains exactly $100r$ per cent of the points of FA.

A mathematical difficulty arises here. It is natural to regard FA geometrically as an uncountably infinite point set. Just how does one identify a subset "containing exactly $100r$ per cent" of these points?" The difficulty can be obviated in an abstract treatment by resort to the Lebesgue concept of area. Below in the text of this paper it will disappear after a simplification of the model.

Purely for notational convenience, it is useful to enlarge the sample space $\{(r, E_r)\}$ by making it consist of all ordered pairs (r, E), where $0 \leq r \leq 1$ and E is any subset of FA [that is, any member of $\mathcal{P}$(FA)]. Of course, for a fixed r, a collection A of subsets E which contains no subsets E of type E_r is logically an "impossible" event. This can be taken care of in the assignment of probabilities by assigning a zero probability to such a collection A.

This very general formulation of the sample space for the precipitation random experiment will now be simplified to accord with the realities of forecasting and of the actual observation of precipitation. The more abstract model and its implications will be studied mathematically in a paper to be published elsewhere.

In the first place, FA, considered as a continuous geometric point set, will be reduced to a finite point set containing N points. These points may be thought of as representing the physical locations of the rain gages in the original FA. But now a revised version of the areal coverage r is required when the observations are confined to these N points. The proportion r henceforth refers to the fraction of the N rain gages which show more than a trace. It can take on only the discrete set of values $r_i = i/N$, $i = 0, 1, 2, \cdots, N$. In the sequel these fractions will often be called the "admissible values" of r.

If the N points in FA are assigned integers from 1 to N, the new sample space can be represented by the set of all pairs $\{r_i, E\}$, where $i = 0, 1, 2, \cdots, N$, and E now is any subset of the set of integers $\{1, 2, \cdots, N\}$. But it is desirable to have a more explicit notation which indicates exactly which of the N points belong to a given set E and which do not. To that end, the points now comprising FA will be numbered from 1 to N, and to the jth point, $j = 1, 2, \cdots, N$, an indexing variable x_j will be assigned which has only two values, 0 or 1. A particular subset E of the N points can now be uniquely represented by an N-tuple $(x_1, x_2 \cdots, x_N)$ in which $x_j = 1$ if the jth point is in E, and $x_j = 0$ otherwise. For example, with $N = 5$, the 5-tuple (1, 1, 0, 1, 0) represents a subset of the 5 points which consists of the first, second and fourth point. The proportion r of the rain gages in the FA contained in a point E can now be expressed as [number of ones in the N-tuple representing E]$/N = (\sum_1^N x_j)/N$. Thus, $\sum_1^N x_j = rN$.

To summarize, in set theoretic notation the simplified sample space is now represented by the set
$\{(r, x_1, x_2, \cdots, x_N):$

$$r = i/N,\ i = 0, 1, \cdots, N;\ x_j = 0, 1, j = 1, \cdots, N\}.$$

(There are $(N+1)2^N$ simple events in this sample space.) For example, with $N = 5$, the symbol (0.4, 0, 1, 0, 1, 0) represents the event that $100r$ per cent ($= 40\%$) of the gages show more than a trace, and the gages numbered 2 and 4 are exactly the ones which do show more than a trace. The symbol (0.2, 0, 1, 0, 1, 0) is an "impossible" event which will be given a probability zero in the assignment of probabilities. The sample space $\{(r, x_1, x_2, \cdots, x_N)\}$ will henceforth be denoted by Ω.

3. Assignment of the probabilities

In view of the intended application of the model, the natural way to assign a probability measure to the sample space Ω is as follows. For a specific admissible value of r, say $r = r_i = i/N$, a (conditional) probability distribution is assigned to the subset of Ω,

$$S_r = \{(r_i, x_1, x_2, \cdots, x_N): x_j = 0, 1, j = 1, \cdots, N\}.$$

This is done for each admissible value of r. The assignment of probabilities is completed by specifying an (unconditional) probability distribution for r on the points $r=r_i=i/N$, $i=0, 1, \cdots, N$.

The probability distribution for r will be represented by the $(N+1)$ vector $\boldsymbol{\pi}=(\pi_0, \pi_1, \cdots, \pi_N)$ where $\pi_i=\text{Prob}(r=r_i)\geqq 0$, $i=0, 1, \cdots, N$, and $\pi_0+\pi_1+\cdots+\pi_N=1$. The number π_0 is equal to the probability of less than a trace in all gages ("no rain") in FA during the forecast period and the number π_N is the probability of more than a trace in all gages ("general rain").

The conditional probability distribution assigned to to S_r for a fixed r will be specified pointwise. The conditional probability given to the point $(r, x_1, x_2, \cdots, x_N)$ will be denoted by $p=p(x_1, x_2, \cdots, x_N|r)$. If $x_1+x_2+\cdots+x_N\neq rN$, then the probability p is assigned the value zero. If $x_1+x_2+\cdots+x_N=rN$, then p may or may not be assigned the value zero, but it must be non-negative, and the following restriction must be observed to accord with the standard axioms of probability theory:

$$\left.\begin{array}{c}\displaystyle\sum_{x_1=0}^{1}\sum_{x_2=0}^{1}\cdots\sum_{x_N=0}^{1} p(x_1, x_2, \cdots, x_N|r)=1 \\ x_1+x_2+\cdots+x_N=rN\end{array}\right\}. \tag{1}$$

The number of terms in this sum is equal to the number of distinguishable ways in which rN ones and $N-rN$ zeros can be arranged in a row, and is given by the formula

$$\binom{N}{rN}=\frac{N!}{(rN)!(N-rN)!}.$$

If $r=0$, then $p(0, 0, \cdots 0|r)=1$ and all the other conditional probabilities are zero. If $r=1$, then $p(1, 1, \cdots, 1|r)=1$ and all the other conditional probabilities are zero.

In the overall sample space Ω, the event that r equals some admissible fixed value, say r_i, and that the Jth rain gage shows more than a trace (so $x_J=1$), is comprised of the set of simple events $\{(r_i, x_1, x_2, \cdots, x_N): x_j=0 \text{ or } 1, j=1, 2, \cdots, J-1, J+1, \cdots, N; x_J=1\}$ with the restriction that $x_1+x_2+\cdots+x_{J-1}+1+x_{J+1}+\cdots+x_N=r_iN=i$. The conditional probability of this event is

$$\left.\begin{array}{c}\displaystyle p(x_J=1|r_i)=\sum_{x_1=0}^{1}\sum_{x_2=0}^{1}\cdots\sum_{x_{J-1}=0}^{1}\sum_{x_{J+1}=0}^{1}\cdots\sum_{x_N=0}^{1} \\ p(x_1, \cdots, x_{J-1}, 1, x_{J+1}, \cdots, x_N|r_i) \\ x_1+x_2+\cdots+x_{J-1}+1+x_{J+1}+\cdots+x_N=i\end{array}\right\}. \tag{2}$$

THEOREM I. *For each admissible value r_i of r, the arithmetic mean of the conditional probabilities $p(x_1=1|r_i)$, $p(x_2=1|r_i)$, $\cdots$, $p(x_N=1|r_i)$, equals r_i. In symbols*

$$\frac{\sum_{j=1}^{N} p(x_j=1|r_i)}{N}=r_i.$$

The proof is given in the Appendix.

Now let $p(x_J=1)$ denote the unconditional probability that the Jth location receives rain. *This is the "point probability" for the Jth location.* By a standard formula[5]

$$p(x_J=1)=\sum_{i=0}^{N}\pi_i p(x_J=1|r_i). \tag{3}$$

(The first term in the sum is zero and the last is π_N.)

THEOREM II. *The arithmetic mean of the point probabilities at the N locations is equal to the theoretical mean value $E(r)$ of the proportion r of area covered.* (This mean value is calculated over the assigned probability distribution of r represented by the vector $\boldsymbol{\pi}$.) In symbols,

$$\frac{\sum_{j=1}^{N} p(x_j=1)}{N}=E(r).$$

For according to Eq. (3) and Theorem I,

$$\frac{\sum_{j=1}^{N} p(x_j=1)}{N}=\sum_{i=0}^{N}\pi_i\left[\frac{\sum_{j=1}^{N} p(x_j=1|r_i)}{N}\right]=\sum_{i=0}^{n} r_i\pi_i.$$

The right member of this equation is by definition the theoretical mean value of r, considered as a random variable with a probability distribution given by the vector $\boldsymbol{\pi}$.

The analogous theorems for multipoint probabilities will now be presented. Let $p(x_{J_1}=1, x_{J_2}=1, \cdots, x_{J_m}=1|r)$ be the conditional probability that the gages at the particular locations $J_1, J_2, \cdots, J_m$, $m\leqq rN$, all show more than a trace given that rN of the rain gages each show more than a trace. Then for $r=r_i$,

$$p(x_{J_1}=1, x_{J_2}=1, \cdots, x_{J_m}=1|r_i)=\sum p(x_1, x_2, \cdots, x_N|r_i), \tag{4}$$

where the summation is over all N-tuples $(x_1, x_2, \cdots, x_N)$ in which $x_{J_1}=x_{J_2}=\cdots=x_{J_m}=1$ and $x_1+x_2+\cdots+x_N=r_iN=i$.

Let $p(x_{J_1}=1, x_{J_2}=1, \cdots, x_{J_m}=1)$ denote the m-point *unconditional* probability that the particular locations $J_1, J_2, \cdots, J_m$ all register rainfall in the forecast period. Then the analogue of (3) is

$$p(x_{J_1}=1, x_{J_2}=1, \cdots, x_{J_m}=1)=\sum_{i=0}^{N}\pi_i p(x_{J_1}=1, \cdots, x_{J_m}=1|r_i). \tag{5}$$

[5] See Feller (1957, p. 106).

The terms in the summation with $r_i < m/N$ are zeros.

It is to be noted that all the terms on the right side of Eq. (4) are included among those on the right side of Eq. (2), so the following inequalities are valid:

$$p(x_{J_k}=1|r_i) \geqq p(x_{J_1}=1, x_{J_2}=1, \cdots, x_{J_m}=1|r_i); \quad k=1, 2, \cdots, m; i=0, 1, 2, \cdots, N. \tag{6}$$

It then follows from comparison of Eq. (5) with Eq. (3) that

$$p(x_{J_k}=1) \geqq p(x_{J_1}=1, \cdots, x_{J_m}=1) \tag{7}$$

with $k=1, 2, \cdots, m$, and for any selection of m locations.

Another obvious inequality which might be worth mentioning is

$$p(x_{J_1}=1, x_{J_2}=1, \cdots, x_{J_m}=1) \geqq \pi_N, \tag{8}$$

where π_N is the probability of 100% areal coverage (a general rain). The inequality is true for $m=1, 2, \cdots, N$. If the forecast is based on the assumption that there are just the two possibilities, "no rain" and "general rain," then the relation (8) is an equality. The validity, in general, of the relation (8) follows from the fact that $p(x_{J_1}=1, \cdots, x_{J_m}=1|r=1)=1$, and the sum in Eq. (5) is at least as large as its last term which is $\pi_N \times 1$.

For each fixed admissible r and $m \leqq rN$, there are $\binom{N}{m}$ conditional probabilities $p(x_{J_1}=1, x_{J_2}=1, \cdots, x_{J_m}=1|r)$, corresponding to the number of different selections of the set of m integers $\{J_1, J_2, \cdots, J_m\}$ from the set of integers $\{1, 2, \cdots, N\}$. The extensions of Theorem I and II to the multipoint case will now be presented.

THEOREM III. *With $m \leqq rN$,*

$$\frac{\sum_{J_1, J_2, \cdots, J_m} p(x_{J_1}=1, x_{J_2}=1, \cdots, x_{J_m}=1|r)}{\binom{N}{m}} = r\frac{(rN-1)(rN-2)\cdots(rN-m+1)}{(N-1)(N-2)\cdots(N-m+1)},$$

where the summation is over all selections of m integers from the first N positive integers.

The proof will be deferred to the Appendix. Then by using Eq. (5) and treating r as a random variable with a distribution given by the vector π, we obtain

THEOREM IV.

$$\frac{\sum_{J_1, J_2, \cdots, J_m} p(x_{J_1}=1, \cdots, x_{J_m}=1)}{\binom{N}{m}} = E[f(r)],$$

where

$$f(r) = r\frac{(rN-1)(rN-2)\cdots(rN-m+1)}{(N-1)(N-2)\cdots(N-m+1)},$$

and the summation is over all selections of m integers from the first N positive integers.

The function $f(r)$ which appears in Theorems III and IV is approximately equal to r^m for large N. It is easy to show algebraically that $f(r) < r^m$ for $0<r<1$.

In fitting this model in full detail to a series of forecasts, it is logical to regard the conditional probabilities $p(x_1, x_2, \cdots, x_N|r)$ as being determined by local climatological conditions. Therefore, they are supposed to be set up numerically once and for all for each admissible value of the proportional areal coverage r. This could be done by using statistical data giving which, and how many, rain gages showed more than a trace in a long series of forecast periods. As a first-order approximation to such a detailed procedure, attention can be concentrated on the conditional one-point probabilities $p(x_J=1|r)$, which are then supposed to be determined (presumably again by statistical methods) *once and for all* for each location J and for each admissible areal coverage r. *It is only the probability distribution of r, given by the vector $\pi=(\pi_0, \pi_1, \cdots, \pi_N)$ which would be the object of estimation in a given forecast.*

The italicized words "once and for all' in the preceding paragraph need qualification. Several sets of values of the conditional probabilities $p(x_1, x_2, \cdots, x_N|r)$ or $p(x_J=1|r)$ may be needed if the local climatological conditions change seasonally, or indeed diurnally. As an example, in a Miami and vicinity forecast, if the forecast period is from 0000 to 1200 GMT and if 20% of the area will be wetted, then at certain times of the year and under certain conditions rainfall occurs only in convective showers which move in from over the Gulf Stream and the 20% will consist of only a narrow costal strip. On the other hand, if the forecast period is from 1200 to 0000 GMT on a summer day, and if 20% of the area will be wetted, often the affected area will be close to the Everglades, well west of the heavily populated districts. What is meant here by "once and for all" is that in this model *the (unconditional) probabilities of the various values of r represented by the vector π are the parameters which vary from one precipitation probability forecast to another* and the values of the probabilities $px_1, x_2, \cdots, x_N|r)$ are to be thought of as predetermined constant coefficients insofar as the forecast is concerned.

4. Mathematical restrictions inherent in the model

In this section, the model developed in the two preceding sections will be examined to see what restrictions the assumption of a single uniformly applicable point probability imposes on the assignment of probabilities to the sample space and on the interpretation of the forecast. A precipitation point probability which is assumed to be valid at all points of FA will be called a uniform point probability. The first finding will be that a uniform point probability is numerically equal to the theoretical mean value of the proportionate areal coverage. In more detail, this theoretical mean value, denoted by $E(r)$, is the mean or "expected" value of the fraction r of the N rain gages which will show more than a trace, when r is treated as a random variable with a probability distribution which is consistent with the point probability forecast. This mean value is not necessarily coincident with any admissible value of r. For example, if the forecast probability is 40% and there are $N=6$ rain gages, then the admissible values of r are 0, 1/6, 2/6, 3/6, 4/6, 5/6, 6/6, and none of these fractions equals 40/100.

Operationally, from a statistician's point of view, $E(r)$ is the arithmetic average of the values of which would be observed in a long sequence of cases in which a particular point probability, like 40%, is correctly forecast.

The fact that a uniform point probability is equal to the theoretical mean value of the proportion of area covered follows at once from Theorem II. Let $p(x_j=1)=P$, a constant, $j=1, 2, \cdots, N$ in that theorem. The equation in the theorem becomes

$$E(r)=\frac{\sum_{j=1}^{N} P}{N}=\frac{NP}{N}=P.$$

Similarly, Theorem IV shows that if the m-point probabilities are uniform, their common value is approximately equal to (but less than) the theoretical mean value of r^m.

The second finding is that a policy of always issuing a uniform point probability for FA, no matter what the estimated probability distribution of the proportionate area coverage r may be, implies that given any admissible r, the corresponding one-point conditional probabilities $p(x_J=1|r)$, $J=1, 2, \cdots, N$, must have a common value. This value is r.

The truth of this proposition depends vitally on persistence on the part of the forecaster in releasing uniform point probabilities in the presence of varying probability distributions of r. It is *not* generally true that if the assignment of the probability to the sample space Ω is such that the unconditional point probabilities $p(x_j=1)$, $J=1, \cdots, N$, happen to be all equal, then necessarily the conditional point probabilities $p(x_j=1|r)$, $J=1, \cdots, N$, are equal for each admissible value of r. In general, there are far too many degrees of freedom in Eqs. (2) and (3) taken together to permit such a conclusion to be drawn. What is true is this: Consider any assignment of the conditional point probabilities $p(x_1, x_2, \cdots, x_N|r_i)$ which yields through Eq. (2) a certain matrix of one point conditional probabilities,

$$n=\begin{bmatrix} p(x_J=1|r_i) \\ J=1, \cdots N \\ i=0, 1, \cdots, N \end{bmatrix}.$$

(The row index is J and the column index is i.) It will be recalled that in the intended application the elements of this matrix are supposed to be determined "once and for all" by local climatological conditions and are independent of the probability distribution assigned to r, which will vary from forecast to forecast. Suppose now that for every assignment of the probability vector π for r, it is known that the unconditional point probabilities given by Eq. (3) are all equal. Then the only specifications of the conditional point probabilities $p(x_1, \cdots, x_N|r)$ which are consistent with this information are those for which M has the form

$$M=\begin{bmatrix} 0 & r_1 & r_2 & \cdots & r_{N-1} & 1 \\ 0 & r_1 & r_2 & \cdots & r_{N-1} & 1 \\ 0 & r_1 & r_2 & \cdots & r_{N-1} & 1 \end{bmatrix}.$$

The proof is exceedingly simple. One of the theoretically possible assignments of the probability distribution of r is that in which for a selected integer I, $0<I\leq N-1$, the only possible values for r are taken to be r_I and 0, the former say with probability q and the latter with probability $1-q$. Then for each J, $J=1, 2, \cdots, N$, Eq. (3) becomes simply

$$\begin{aligned} p(x_J=1) &= (1-q)p(x_J=1|0)+q\cdot p(x_J=1|r_J), \\ &= (1-q)\cdot 0+q\cdot p(x_J=1|r_I), \\ &= q\cdot p(x_J=1|r_I). \end{aligned}$$

But, by hypothesis, there exists a constant P such that $p(x_J=1)=P$, $J=1, \cdots, N$. Therefore, $p(x_J=1|r_I)=P/q$, $J=1, \cdots, N$. By Theorem I, the only possibility now is that $p(x_J=1|r_I)=r_I$. The argument can be repeated for each integer I, $0<I<N$. Incidentally, the elements in the first and last columns of M are automatically zeros and ones, respectively, by definition.

The result can be stated formally as follows:

THEOREM V. *If no matter what probability distribution is assigned to r (the proportion of rain gages in FA showing more than a trace), the unconditional point probabilities satisfy the relation*

$$p(x_1=1)=p(x_2=1)=\cdots=p(x_N=1),$$

then the only assignments of the areal conditional probabilities $p(x_1, x_2, \cdots, x_N|r)$ *which are compatible with this assumption are those in which the conditional one-point probabilities obey the relations*

$$\begin{aligned} r_i=p(x_1=1|r_i)&=p(x_2=1|r_i) \\ &=\cdots=p(x_N=1|r_i),\ i=0, 1, \cdots, N. \end{aligned}$$

The proof given above depends on the assumption of persistence of the equality of the point probabilities even in the presence of certain rather extreme specifications of the probability distribution of r. The next theorem shows that the conclusions of the theorem remain true if equality of the point probabilities is assumed for only a few reasonable specifications of the vector π which gives the probability distribution of r.

THEOREM VI. *Let S be a set of N specifications of the vector π of which the kth is denoted by*

$$\pi^k = (\pi_0{}^{(k)}, \pi_1{}^{(k)}, \cdots, \pi_N{}^{(k)}); k=1, \cdots, N.$$

Further let S be such that the N-dimensional vectors $(\pi_1{}^{(k)}, \pi_2{}^{(k)}, \cdots, \pi_N{}^{(k)})$, $k=1, 2, \cdots, N$, are linearly independent. If it is assumed that the point probabilities $p(x_J=1)$, $J=1, 2, \cdots, N$, are equal for each vector π^k, then

$$r_i = p(x_1=1|r_i) = \cdots = p(x_N=1|r_i),\ i=0, 1, \cdots, N.$$

The common value of the point probabilities for a given vector π^k depends on k. The proof of the theorem is given in the Appendix.

The special specifications of the vector π used in the proof of Theorem V form a set S of the type described in Theorem VI. *In fact, any set of N specifications of π in which the kth vector π^k gives preferential probability to the value $r=r_k=k/N$, over all alternative intermediate proportions of area covered for each k, $k=1, 2, \cdots, N$, satisfies the requirements on S in the Theorem.* (This is proved in the Appendix.) The words "preferential probability" are to be interpreted by the following inequality among the components of π^k:

$$\pi_k{}^{(k)} > \pi_1{}^{(k)} + \pi_2{}^{(k)} + \cdots + \pi_{k-1}{}^{(k)} + \pi_{k+1}{}^{(k)} + \cdots \pi_N{}^{(k)}.$$

This seems to be a natural type of specification.

For example, it would include something of this sort. With $N=10$, the probability of 60% coverage is 0.7 of 50% or 0.2, of 70% is 0.1 and the probability of any other coverage is zero. The vector π in this example is $(0, 0, \cdots, 0.2, 0.7, 0.1, 0, \cdots, 0)$.

Qualitatively speaking, the force of Theorem VI is this: If the model developed in Sections 2 and 3 is used to prepare and interpret precipitation probability forecasts, and if they are repeatedly issued as a single point probability assumed to be applicable to the entire FA, then there is an assumption that for any intermediate areal coverage, the conditional probability that a rain gage shows more than a trace is the same for all gages in FA.

Uniform conditional point probabilities, given a fixed proportion r of area coverage, do *not* necessarily imply that the precipitation is "selecting at random" the rN gages which show more than a trace. (Here "selection at random" means that the conditional probability attached to any particular set of rN gages is the same for all sets of rN gages.) Given any set of N numerical conditional one-point probabilities, whether all the same or not, Eqs. (2) allow wide latitude in the corresponding conditional probabilities $p(x_1, x_2, \cdots, x_N|r)$ even for quite small values of N, because as equations in these unknowns they are in general under-determined. These are, in fact, N equations in $N!/[(rN)!(N-rN)!]$ unknowns.

But if it is assumed that precipitation covering a proportion r of the gages does select the gages at random in the above sense for each value of r, then this implies that the conditional point probabilities $p(x_1, x_2, \cdots, x_N|r_i)$ for $r=r_i$ are all equal. $\left[\text{The common value would be } 1\Big/\binom{N}{i}.\right]$ It would follow from Eqs. (2) and (4) that the conditional one-point and conditional multipoint probabilities would all be equal, and then according to Eqs. (3) and (5), the unconditional one-point and multipoint probabilities would have to be equal.

With the methods used to establish Theorem VI it can be shown that if the possible specifications of the probability distribution of r include those of the type described in the hypotheses of this theorem, then an assumption of uniform m-point (unconditional) probabilities, $m>1$, implies that the conditional m-point probabilities must be all equal for each admissible value of r, whether or not random selection of gages is assumed. When such is the case, then according to Theorem III the common value of these conditional probabilities for a given r must be the right member of the equation in Theorem III. In particular, the 2-point probabilities must all have the value $r(rN-1)/(N-1)<r^2$.

It seems apparent from the following argument that in actual practice the assumption of uniform 2-point probabilities would not be acceptable when uniform 1-point probabilities are also assumed. If the conditional 1-point probabilities are assumed to be all equal for a given value of r, then according to Theorem I, each must be numerically equal to r itself. Consider now the following two events: 1) the gage at location J shows more than a trace; and 2) the gage at location K shows more than a trace. For the given value of r, if these were statistically independent events, then the probability that they *both* occur during the forecast period would be the product of the two point probabilities, or r^2. But if the assumption of uniform 2-point probabilities is made, then as seen in the preceding paragraph, the probability that the two events both occur is less than r^2; in other words, it is even less than it would be with statistical independence. Yet if the locations J and K were relatively close together geographically, one would expect on practical grounds that the probability that both gages show rain would be nearly equal to the probability that either one shows rain, regardless of what happens in the other gage. This probability, under the assumption of a uniform one-point probability is equal to r, and r is greater than r^2 when r is not 0 or 1. Thus, a uniform 2-point probability seems to be incompatible with a uniform 1-point probability.

5. Precipitation probability forecasting in practice as related to the model

There are probably a number of different methods by which U. S. Weather Bureau forecasters are arriving at precipitation probability forecasts. However, the one about to be described has been officially sanctioned,[6] and it is the only one which will be discussed in this paper.

Assume that there is only a finite set of distinguishable circulation regimes $R_1, R_2, \cdots, R_K, \cdots, R_n$, which produce precipitation over FA, and when such a regime is present during the forcast period, some rain will surely fall somewhere in FA. Assume further that for each regime R_K there exists a uniquely determined probability for each positive admissible value of r (the proportion of area covered). Let Prob $(r=r_i)=f_i^{(K)}$, $i=1, 2, \cdots, N$, where $r_i=i/N$, $i=0, 1, \cdots, N$, are the admissible values of r, and $f_1^{(K)}+f_2^{(K)}+\cdots+f_N^{(K)}=1$. In the framework of the model of Sections 2 and 3, the vector $\mathbf{f}^K=(f_1^{(K)}, \cdots, f_N^{(K)})$ plays the role of a conditional probability distribution of r, given that R_K has occurred. Suppose that by data processing of past observations on the variable r for the various regimes R_K, an arithmetic average A_K of the observed values of r has been computed for each R_K. For each given R_K, A_K is theoretically an optimal estimate of the conditional mean value

$$E_K(r|R_K)=\sum_{i=1}^{N} r_i f_i^{(K)}.$$

The forecaster's decision process is then this: At forecast time he identifies a particular regime R_K which seems likely to obtain over FA during the forecast period. He then proceeds on the assumption that either R_K will indeed obtain during the forecast period or there will be no rain at all anywhere in FA. He then estimates the probability Prob $(r>0)$, of some rain somewhere [or the complementary probability $\text{Prob}(r=0)$], multiplies this by A_K, and announces the product as the forecast point precipitation probability.

Disregarding statistical approximations, the number P thus arrived at is given by the formula

$$P=E(r|R_K)\cdot\text{Prob}(r>0). \tag{9}$$

Now in what sense is this a point probability valid for the entire FA?

The selection of R_K and the probability $\text{Prob}(r>0)$ of its presence selects automatically an unconditional probability distribution $\boldsymbol{\pi}^K=(\pi_0^{(K)}, \pi_1^{(K)}, \cdots, \pi_N^{(K)})$ for r over the entire range of admissible values of r including $r_0=0$. The components of $\boldsymbol{\pi}^K$ are, respectively, $\pi_0^{(K)}=\text{Prob}(r=0)$, $\pi_i^{(K)}$

$$=f_i^{(K)}\cdot\text{Prob}(r>0), \quad i=1, 2, \cdots, N.$$

(It will be noted that the sum of these components is unity, as it should be.) Then the term on the right side of Eq. (9) can be transformed as follows:

$$E(r|R_K)\ \text{Prob}(r>0)=(\sum_1^N r_i f_i^{(K)})\cdot\text{Prob}(r>0)$$

$$=\sum_{i=1}^{N} r_i\pi_i^{(K)}=\sum_{i=0}^{N} r_i\pi_i^{(K)}=E_K(r), \tag{10}$$

where $E_K(r)$ is the unconditional mean value of r over the probability distribution given by $\boldsymbol{\pi}^K$. So here the vector $\boldsymbol{\pi}^K$ has been specified partly by choosing a regime R_K which automatically specifies the numerical probability vector $\mathbf{f}^K$, and partly by choosing a numerical value for the probability $\text{Prob}(r>0)$ that this regime R_K actually will obtain over FA during the forecast period. Thus, when a forecaster designates the right member of Eq. (9) as a point probability, he is actually saying that the point probability is the mean value of areal coverage $E_K(r)$.

To justify this rigorously, it is apparently necessary to refer to the probability model developed in Sections 2 and 3. It is brought out in Section 4 that within this model, a uniform point probability P does imply an assignment of probabilities to the sample space which is such that the mean value of the proportion r of area covered is equal to P.

But the assumption of a uniform point probability is not without implications as to the conditional point probabilities, and these implications may be incompatible with physical realities. The implications are contained in Theorem VI of Section 4, which will now be interpreted in the language of the present discussion. Suppose that the set of circulation regimes $R_1, R_2, \cdots, R_n$ which a forecaster recognizes in making forecasts contains a subset of N regimes, say $R_{K1}, R_{K2}, \cdots, R_{KN}$, for which the corresponding probability vectors $\boldsymbol{\pi}^{K1}, \boldsymbol{\pi}^{K2}, \cdots, \boldsymbol{\pi}^{KN}$ are a linearly independent set of vectors. Suppose further that the forecaster always issues a uniform point probability or at least always does so when confronted with any of the regimes $R_{K1}, \cdots, R_{KN}$. Then according to Theorem VI, the forecaster must be assuming that the conditional one-point probabilities, given any particular proportion r of area coverage, must all be equal (and equal to r).[7] For example, given that

[6] See Hughes (*op. cit.*), Eq. (1), and accompanying explanation. The description given in the text above is a paraphrase in the terminology of this paper of information transmitted to the author by C. F. Roberts, Chief, Technical Procedures Branch, U. S. Weather Bureau, in a private communication.

[7] The hypothesis of Theorem VI requires the existence of a set of specifications $\boldsymbol{\pi}^k$, $k=1, \cdots, N$, of the unconditional probability vector $\boldsymbol{\pi}$ in which the N vectors formed by striking out the first component of each $\boldsymbol{\pi}^k$ are linearly independent. It is indicated in the discussion following Eq. (9) that when the forecaster selects the regime R_{Kk} and a probability Prob_F $(r>0)$ that R_{Kk} will obtain, where Prob_F $(r>0)$ varies with the forecast, then he is in effect constructing an unconditional probability vector $\boldsymbol{\pi}^k$ for r with components $[\text{Prob}_F\ (r=0),\ f_i^{(Kk)}\ \text{Prob}_F\ (r>0),\ i=1, \cdots, N]$. If the N vectors $\mathbf{f}^{Kk}$, $k=1, \cdots, N$, are linearly independent, then the N vectors formed, respectively, by taking the dot products of the vectors $\mathbf{f}^{Kk}$ with any set of N positive scalars Prob_{F1} $(r>0)$, Prob_{F2} $(r>0), \cdots, \text{Prob}_{FN}$ $(r>0)$ will again be linearly independent. Thus, the admission of the regimes $R_{K1}, R_{K2}, \cdots, R_{KN}$ into the repertory of the forecaster implies the existence of a set of specifications of the unconditional probability vector $\boldsymbol{\pi}$ as required in the hypothesis of Theorem VI.

there was only 10% coverage, the probability that any gage showed more than a trace would be the same (10%) for all gages in FA.

It would be out of place in this paper to discuss at length the compatibility of such a restriction with physical reality. In the case of certain local forecasts for Miami and vicinity mentioned at the end of Section 3, it does seem rather meaningless at times to release a single point probability. It also seems to the author to be confusing to issue a uniform point probability, like 50%, and then qualify it by words such as "well inland during afternoons," as is done in the forecast quoted from the *Miami Herald* in the Introduction. In this example, what does that 50% really mean? Apparently the 50% is not even valid at the official verification gage at the Miami International Airport, because this is not "well inland."

Three rather obvious remedies come immediately to mind for cases in which a uniform point probability is suspect. The first one can be criticized as merely a sop to the purist, but at least the forecaster would be making a theoretically correct statement. This remedy consists in identifying a precipitation probability forecast as an arithmetic average of the true point probability values at the points of FA. The actual words used would be something like this: "Shower probability, 50 per cent on the average for the area," or "Average shower probability, over the area, 50 per cent." It would not be contradictory to qualify this kind of statement by identifying subregions of FA where higher or lower point probabilities are expected to prevail.

The other two suggestions involve issuing different point probabilities for different subregions of FA. First, there is the obvious procedure of making separate forecasts for various subregions by the methods used for the overall forecast. It would appear that this might impose an unjustifiable additional burden on the local forecast stations.

The second suggestion involves exploiting for practical purposes the model in this paper. Just what complications would that entail?

Assuming that the forecaster takes the approach described earlier in this section, for each R_K he would need to have approximations (based on historical experience) for the corresponding individual probabilities $f_1^{(K)}, f_2^{(K)}, \cdots, f_N^{(K)}$ instead of merely an approximation to the theoretical mean value $E(r|R_K)$ of this distribution. If the historical average value of r has been actually computed for a given R_K and not just guessed at, then the data on which it was computed might also be available, and these data consist of the relative frequency approximations to the probabilities $f_i^{(K)}$.

In the second place it would be necessary to have estimates of the conditional one-point probabilities $p(x_J=1|r)$ for the various admissible values of r at the N locations of rain gages in FA. The datum for a determination of the value of $p(x_J=1|r)$ for a fixed J and for a given admissible value of r, say r_i, consists merely of the relative frequency with which more than a trace of precipitation was observed in the Jth gage in a suitably large number of observed forecast periods in which a proportion r_i of the N gages showed rain. Although these conditional one-point probabilities are supposed to be functions of local climatological conditions and independent of the probability distribution of r, nevertheless if the method of forecasting by selection of a likely regime R_K is used, it may be desirable to have two or more different sets of estimates available for the conditional one-point probabilities, corresponding, respectively, to two or more different classes of circulation patterns R_K.

The forecaster who uses the approach described at the beginning of this section would now replace Eq. (9) in calculating the probability forecast by the following equation, which is a restatement of Eq. (3) in the notation used in this section:

$$p(x_J=1)=\mathrm{Prob}(r=0)+\mathrm{Prob}(r>0)\cdot\sum_{i=1}^{N} f_i^{(K)}p(x_J=1|r_i). \quad (11)$$

Although a different point probability would be calculated for each point J, presumably only two or three approximations relevant to certain designated subregions of FA would be released.

If approximations to the detailed conditional probabilities $p(x_1, x_2, \cdots, x_n/r)$ were to be available instead of only the one-point conditional probabilities $p(x_J=/r)$, then a forecast of the probability vector $(f_1^{(K)}, \cdots, f_n^{(K)})$ and of Prob $(r>0)$ could be used to calculate probabilities of various events concerning which the one-point probabilities give no information except through inequalities such as (6), (7) and (8). For example, with additional information, the multipoint unconditional probabilities appearing in Eq. (5) could be easily calculated.

If the two-point unconditional probabilities were to be calculated for two specific points numbered J and K in FA, then a problem such as "What is the probability that J gets rain and K does not" can be solved. [The answer to this one is $p(x_J=1)-p(x_J=1,\ x_K=1)$.] Again, "What is the probability that neither one of the two locations J and K gets rain?" [Answer, $1-p(x_J=1)-p(x_K=1)+p(x_J=1,\ x_K=1)$.]

6. Some fundamental questions regarding precipitation probabilities

There apparently remains a certain amount of controversy about the usefulness of releasing numerical precipitation probabilities to the public, and also about the significance (or lack of it) of the released numbers as true probabilities. It seems worth while to identify here two areas in which questions might arise. No answers will be given, and no adverse criticism of U. S. Weather Bureau policy as to precipitation probability forecasting is intended. Rather, the idea is to suggest new types of statistical verification programs which might clarify the situation.

The discussion in this section is concerned only with the sequence of rain-no rain observations taken at the official verification station of a local forecast region in relation to the announced "probability numbers." There will be no further consideration of the areal distribution problem which forms the main theme of preceding sections. Attention will be centered chiefly on two topics: (a) Assuming that precipitation probability forecasts are possible within the framework of classical probability theory, have systematic biases in estimation been visible in practice, which may be due to a scoring method? (b) Do sequences of observed rain-no rain observations for given forecast "probability" categories exhibit a sufficient degree of randomness to justify treating the released "probability numbers" as probabilities?

A discussion of the question of randomness in numerical sequences may be found, for example, in Feller (1957, Chap. VIII). For present purposes, the sequences can be taken to be sequences of observations on a dichotomous variable X which assume only the values 0 and 1. It is not useful to try to define "randomness" by algebraic means for a single specific sequence, such as 0110111010. The usual modern approach is via testing the hypothesis that an observed sequence is generated by successive independent observations on a random variable X which has a fixed basic probability for the event $X=1$ in each trial. This hypothesis will be called the hypothesis of randomness. The sample space for the hypothesized random experiment consists of all possible infinite sequences of zeros and ones. It is not difficult to calculate under the hypothesis of randomness the probability of many different kinds of events in this sample space. For example, consider the event E consisting of all sequences in which the pattern 010 appears only a finite number of times. The probability of E happens to be zero, and if the event E were visible in a given sequence there would be doubt as to the validity of the hypothesis of randomness. A more crucial consequence of the hypothesis is that if any rule is set up which selects an infinite subsequence of a typical sequence $X_1, X_2, X_3, \cdots$ of determinations of X by accepting or rejecting $X_j, j=1, 2, \cdots$, without regard to the value of X_j (e.g., "reject every other X_j"; or "accept or reject X_j on the flip of a coin"), then the frequency ratio of ones in the first m terms of the subsequence converges to the basic probability P with probability one as m becomes infinite. This means that the event consisting of all sequences in the sample space for which this convergence takes place has probability measure one. (The rule in which every determination of X is accepted is of course included.) The event would change if the rule is changed, but it still would have probability one. The repeating sequence 0000000111000000111$\cdots$, in which the frequency ratio of ones converges to 0.3, violates the subsequence consequence of the hypothesis of randomness in countless ways. If this sequence were observed, and then the repeating sequence 1011000000101100000$\cdots$ (with limiting frequency ratio again 0.3) is next observed, the hypothesis of a random mechanism producing these sequences with basic probability $p=0.3$ would be pragmatically untenable.

The question of bias will now be discussed. There is ample evidence that skillful forecasters can come up consistently with precipitation "probability numbers" which do exhibit *some* of the statistical attributes of true probabilities when verified at the official rain gage. Verification programs leading to this conclusion have been described by Root (1962), Sanders (1963) and various others.[8] The forecast probabilities have been tested mainly for the presence of two attributes, identified in the literature as *reliability and resolution*. Reliability refers to the degree to which any particular forecast point probability P (e.g., $P=30\%$) is reproduced by the relative frequency of forecast periods in which rain is observed at the official gage, in a series of forecast periods for which this P was released. Resolution refers to the extent to which the frequency distribution of released values of P is dispersed about the climatological relative frequency of rain occurrence at the official gage. Resolution can be evaluated qualitatively by just looking at a histogram of this frequency distribution [as in Root (1962, Fig. 1)], or by more sophisticated methods, of which the most popular now seems to be the formula proposed by Brier (1950). Sanders (1963) showed how to break down the Brier score into two additive components, of which the first measures reliability and the second measures resolution. In its role as a measure of resolution, the Brier score rewards a forecaster who consistently releases low precipitation probability numbers for forecast periods during which no more than a trace is actually observed at the official gage, and who consistently releases high probability numbers for forecast periods when more than a trace is actually observed.

In the verification data which this author has seen, two phenomena are almost always present.[9] The first one is a tendency of forecasters to underestimate the probability of rain (as verified by relative frequency) in the probability categories below the climatological relative frequency and to overestimate the probability (as verified by relative frequency) above the climatological relative frequency. The second phenomenon is the avoidance of forecasts of probability numbers close to the climatological relative frequency, as evidenced by a

[8] See for example Hughes (*op. cit.*) and Dickey (1965).

[9] This includes thirteen months of experience with two verified forecast periods daily for the Miami and vicinity forecasts. The data were prepared for the author by Mr. R. C. Sheets with the permission of Dr. Gordon Dunn.

relative minimum near this point in the height of the histogram representing the frequency distribution of released probabilities.[10] It is true that it is easy to imagine a geographical region (such as a desert) in which a long-range climatological relative frequency of precipitation is generally irrelevant to daily forecasts, and the forecasters would have good reason seldom to name this number in probability forecasts. But when biases such as the two indicated above appear so *consistently* in data from such diverse sources, is it not possible to ask whether they might be induced by playing a scoring system to the detriment of realistic probability forecasting?

The more important and more delicate question of randomness in observed rain-no rain sequences will now be considered. In this discussion, the digit 1 will designate the event that rain is observed at the official verification gage, and the digit 0 will denote no rain recorded in that gage.

To lead up to the question let it first be noted that the words "climatological probability" appear frequently in the literature. For example, Sanders (1963) used the term and observed that "a forecast of the climatological probability would be highly valid in the long run" (*loc. cit.*, p. 195). He was making the point that reliability in itself is no measure of forecasting skill, but the issue here is whether a series of precipitation "probability" forecasts consisting of repeated announcements of the climatiological relative frequency would mean anything at all from a probabilistic point of view. Consider the case of the three winter months in San Francisco studied by Root (1962). Rain there during the winter tends to come in periods several days in length, followed by dry spells. The climatological relative frequency is about 30%. But a typical sequence of coded observations covering the three months might read like this: 00···011···100···011···100···0, etc., in which the ratio of "ones" to the total number of symbols is roughly 0.3. The idea is that the "ones" appear in groups and so do the "zeros." If today were the first day of a dry spell, then the precipitation probability for tomorrow is verifiable as 0, not 0.30. The climatological relative frequency of 0.30 cannot be accepted as a forecast probability with day-to-day operational meaning, because the sequence of precipitation events fails quite completely to accord with the hypothesis of randomness.

But let it be taken for granted that "climatological probability" is merely a figure of speech, and that the forecaster, perhaps under the influence of the Brier score, is going to make an effort to sort forecast circulation types into reasonably sharp probability categories. Consider again the data given by Root (1962). It appears, for example, that there were 440 cases in which the forecast probability was 20%, and in these cases, rain was observed 123 times, for a relative frequency of 28.0%. Each time the forecaster issued this forecast, he was predicting the probability that a rainy spell would clear up or a dry spell would be interrupted. It seems possible with such a systematic pattern in the *overall* sequence of observations that the sequence of rain observations in the 440 cases itself might tend to follow some systematic pattern. If as an extreme case they exhibited a pattern such as 0000100001···, then the forecast probability of 20% would be meaningless in making day-to-day interpretations.

The author does not know whether such patterns really did exist in the San Francisco program. The illustration is used only to raise in a specific context the spectra of non-randomness in a sequence of dichotomous precipitation events to which a fixed probability P has been assigned. The author is not aware of any statistical studies which have been made in this direction. There is some reason to suspect trouble. Root (1962) seems to mention obliquely the possibility of lack of independence in precipitation observations for a given probability category. In the Miami data the author found a number of statistical anomalies. One of them was that for the 20% category (23 cases), 50% category (37 cases), 60% category (37 cases), and 70% category (28 cases), the observed relative frequencies of rain deviated by an improbably *small* amount from the forecast probability. This can be symptomatic of non-randomness.

A precipitation "probability number" which is verifiable only as a relative frequency over a long series of forecasts would be economically useful for certain purposes, whether or not the precipitation events appear to be compatible with the hypothesis of randomness. But if the sequence of precipitation events is non-random, the question then arises as to whether some improvement can be made in the relevance of the forecast to individual forecast periods. It may be possible to put the question back into a probability framework by postulating that the overall sequence of rain-no rain observations at the official gage are observations on a two-valued (zero-one) stochastic process (Parzen, 1962, p. 35 ff). A precipitation probability forecast would be an estimate of a transition probability in the process. By transition probability is meant the conditional probability of a "one," given the forecast conditions *and* the values of the observations over the more or less immediate past.

7. Summary and conclusions

This paper is concerned with the precipitation probabilities now issued routinely in local forecasts by the U. S. Weather Bureau. It is assumed that, in general, the occurrence of the event to which a precipitation forecast refers is verifiable by the observation of more than a trace of rain during the forecast period at each point of a particular subregion of the forecast area. The

[10] Both phenomena are present in the data of Root (1962), Hughes (*op. cit.*), Dickey (*op. cit.*, excluding data shown for a Hartford, Conn., program), and the Maimi data referred to in footnote 9.

event to which a precipitation probability forecast applies is the occurrence of more than a trace of rain at a particular point of the forecast area. This event occurs when the point is contained in the wetted subregion.

Under these premises, the logical sample space to use when precipitation events are to be assigned probabilities consists simply of the collection of all subregions of the forecast area. A subregion "occurs" if all points in it receive more than a trace and all other points in the forecast area receive no more than a trace.

To accord with reality, this sample space is simplified by replacing the forecast area by a finite number N of distinct points representing the locations of rain gages in the forecast area. The sample space then becomes the set of all possible selections from these N points, taken one at a time, two at a time, $\cdots$, N at a time. An explicit notation is introduced to identify individual selections by setting up indicator N-tuples $(x_1, x_2, \cdots, x_N)$ in which the components x_J have the values zero and one. If the Jth point is in a given selection, then $x_J=1$ in the N-tuple corresponding to this selection; if not, then $x_J=0$. The sample space can then be represented by the set of all N-tuples $(x_1, x_2, \cdots, x_N)$ in which the elements are zeros and ones.

The assignment of probabilities to this sample space is broken down into two steps in accordance with the intended application of the model. In the first step, the N-tuples are stratified according to the number of ones in each. Let i be the number of ones in a given N-tuple. The fraction $r_i=i/N$ then represents the proportionate areal coverage. For a fixed i, the corresponding N-tuples are each assigned point probabilities $p(x_1, x_2, \cdots, x_N|r_i)$, which add up to unity, and which in the overall assignment of probabilities appear as conditional probabilities, given i(or r_i). The second step consists in assigning an unconditional distribution to a random variable r on the fractions $r_i=i/N$, $i=0, 1, \cdots, N$. In the intended application, the conditional probabilities $p(x_1, x_2, \cdots, x_N|r_i)$ are viewed as constants determined by local climatological conditions, and it is only the distribution of r which is the object of a forecast.

The one-point conditional probability $p(x_J=1|r_i)$ for a fixed areal coverage r_i is given by a suitably restricted summation [Eq. (2)] of the conditional probabilities $p(x_1, x_2, \cdots, x_N|r_i)$ over the sample space. The arithmetic mean of the numbers $p(x_J=1|r_i)$, $J=1, \cdots, N$, is equal to r_i (Theorem I). The unconditional point probability $p(x_J=1)$ of the event that the Jth gage records more than a trace is given by Eq. (3) in terms of the conditional point probabilities $p(x_J=1|r_i)$ and the probability vector $\boldsymbol{\pi}$ or r. The arithmetic mean of the unconditional probabilities for the N locations is equal to the theoretical mean value of r over the distribution given by $\boldsymbol{\pi}$ (Theorem II). Analogous theorems are given for multipoint probabilities (Theorems III and IV). An immediate consequence of the "arithmetic mean" statements in the preceding paragraph is that if for any given r_i, the conditional one-point probabilities $p(x_J=1|r_i)$ are assumed to be equal for $J=1, 2, \cdots, N$, then the common value is r_i, and if the unconditional one-point probabilities $p(x_J=1)$ are assumed to be equal for $J=1, 2, \cdots, N$ (as is usually implied in U. S. Weather Bureau forecasts), then the common value is the theoretical mean value of r over its assigned probability distribution as given by $\boldsymbol{\pi}$. It is proved (Theorem V) that if a forecaster always assumes that the unconditional point probabilities $p(x_J=1)$ are equal, no matter what the forecast vector $\boldsymbol{\pi}$ is, then he is automatically assuming that the conditional point probabilities $p(x_J=1|r_i)$ are all equal for each r_i (and equal to r_i). The word "always" in the preceding sentence can be replaced by merely "sometimes" according to Theorem VI.

An assignment of probabilities leading to equal conditional one-point probabilities $p(x_J=1|r_i)$, $J=1, 2, \cdots, N$, for some particular r_i does not imply that the i gages in which more than a trace is recorded are "selected at random" by the rainfall, in the classical sense that the possible selections of i gages from the N gages are equiprobable.

On both practical and theoretical grounds, an assumption of a uniform two-point probability seems to be inacceptable, especially if at the same time an assumption of a uniform one-point probability is also made. If it is assumed that in cases of partial areal coverage, the gages recording rain are "selected at random," then the two assumptions in the preceding sentence would automatically be implied, so the assumption of random selection should be avoided.

A method of preparing precipitation probability forecasts which is considered by forecasters to be a natural one consists essentially in forecasting the probability vector $\boldsymbol{\pi}$ of r in two steps as follows: First a circulation regime R is selected which seems likely to obtain over FA during the forecast period. If R obtains, there will be some rain somewhere in FA. It is assumed that either R will obtain or there will be no rain. It is presupposed that from past experience the conditional mean value $E(r|R)$ of the proportionate areal coverage, given that R obtains, has been estimated. The forecaster estimates the probability that R will obtain, which is $\mathrm{Prob}(r>0)$. He then releases $E(r|R)\cdot\mathrm{Prob}(r>0)$ as the precipitation probability. It can easily be shown mathematically that the number he is issuing is the theoretical mean value of r over a probability distribution compatible with his selections of R and $\mathrm{Prob}(r>0)$. Then as stated above in this summary, since he is assuming tacitly that his single precipitation probability is valid at each point of the forecast area, it is mathematically true (with this tacit assumption) that the precipitation probability arrived at by the above method is the uniform point probability.

If the repertory of regimes R recognized by a forecaster is as large in number or larger than the number N

of rain gages in the forecast area, and if the repertory contains a subset S of which the probability vectors satisfy the hypothesis of Theorem VI (which seems certain to be the case), then the persistent issuing of a single precipitation probability for the forecast area when confronted by each of the regimes in S implies that the conditional one-point probabilities $p(x_J=1|r)$, $J=1, \cdots, N$, are all equal for each admissible value of r. For certain forecast areas and in certain seasons, such an implication seems to be unrealistic.

In such situations, one rather artificial way to achieve realism is by simply qualifying the released point probability by the word "average," meaning arithmetic average over the forecast area. A precipitation probability calculated by some method which estimates the mean value of the proportionate areal coverage is truly the arithmetic average of the point probabilities at the gages in the forecast area.

Another way is to use the model in this paper to calculate more than one point probability from a single forecast regime R and a forecast of its probability of occurrence. To do this, the forecaster would: (a) need to have an estimate, based on historical data, of the conditional probability distribution of r, given R (instead of merely the conditional mean value of r); and (b) would need to have statistical estimates of the conditional point probabilities $p(x_J=1|r)$, $J=1, 2, \cdots, N$, for the admissible values of r. It would appear that such statistical estimates could very easily be obtained, and that very possibly there already exist historical data in the various forecast regions for making the estimate.

If estimates of the detailed point conditional probabilities $p(x_1, x_2, \cdots, x_N|r)$ are obtained, or at least if estimates of the multipoint conditional probabilities $p(x_{J_1}=1, x_{J_2}=1, \cdots, x_{J_m}=1|r)$ are available, then various multipoint probability problems can be worked out which are inaccessible if only one-point probabilities are given, whether uniform or not. An example is, "What is the probability that neither one of locations J and K gets rain?"

Finally, two questions about precipitation probabilities are raised as typical of areas in which further statistical research might be useful. Both are concerned only with verification at a single official rain gage without regard to areal distribution. The first relates to persistent biases in the direction of "overforecasting" which are visible in the data of almost all published verification programs. The second, a more difficult question, is concerned with the presence or absence of randomness in sequences of rain occurrences within forecast probability categories. Absence of randomness would seriously impair the significance of the released "probability numbers" as actual probabilities.

APPENDIX

Proof of Theorem I

Write our Eq. (2) for each J, $J=1, 2, \cdots, N$, and look at the formula for $p(x_1=1|r_i)$. Consider a particular symbol which appears in it, say $p(1, x_2, x_3, \cdots, x_N|r_i)$. In this symbol, $r_iN-1=i-1$ of the variables $x_2, x_3, \cdots, x_N$ have the value 1. Suppose these are $x_{j_1}, x_{j_2}, \cdots, x_{j_m}$, $m=r_iN-1$. The symbol appears again in the formula for $p(x_{j_1}=1|r_i)$, also in the formulas for $p(x_{j_2}=1|r_i)$, $p(x_{j_3}=1|r_i)$, $\cdots$, $p(x_{j_m}=1|r_i)$. These are the only appearances of this symbol in the set of N formulas for $p(x_J=1|r_i)$, $J=1, \cdots, N$. Similarly, every other symbol in the formula for $p(x_1=1|r_i)$ reappears in exactly r_iN-1 other formulas. By applying this argument to the formulas for $p(x_2=1|r_i)$, $p(x_3=1|r_i)$, $\cdots$, $p(x_N=1|r_i)$, it can be seen that every possible symbol $p(x_1, x_2, \cdots, x_N|r_i)$ appears in the formulas exactly r_iN times. Now sum the formulas. By Eq. (1), the result is

$$\sum_{J=1}^{N} p(x_J=1|r_i)=r_iN\times 1,$$

so

$$\frac{\sum_{J=1}^{N} p(x_J=1|r_i)}{N}=r_i.$$

Proof of Theorem III

For a particular selection of integers $J_1, J_2, \cdots, J_m$, consider a term $p(x_1, x_2, \cdots, x_N|r_i)$ in which (of course) $x_{J_1}=x_{J_2}=\cdots=x_{J_m}=1$, and in which also $x_{J'_1}=x_{J'_2}=\cdots=x_{J'_k}=1$, where $m+k=r_iN=i$, and the subscripts are all different integers. This term will reappear in the formula for $p(x_{K_1}=1, x_{K_2}=1, \cdots, x_{K_m}=1|r_i)$ whenever $\{K_1, K_2, \cdots, K_m\}$ is a selection of m integers from the set of integers $\{J_1, J_2, \cdots, J_m, J'_1, J'_2, \cdots, J'_K\}$. The total number of such selections (including $\{J_1, \cdots, J_m\}$ itself) is $\binom{i}{m}$, so each term on the right side of (4) appears this many times when the Eqs. (4) are written out for all the different selections of m locations from N. Now all possible different N-tuples $(x_1, x_2, \cdots, x_N)$ with exactly i ones in them appear in the Eqs. (4) when they are written out for all the different selections of m locations from N, and by (1), the probabilities $p(x_1, x_2, \cdots. x_N|r_i)$ must add up to one. Therefore, if the $\binom{N}{m}$ equations (4) for a fixed m and r_i are all added together, the right side will equal 1 added to itself $\binom{i}{m}$ times. The resulting equation is

$$\sum p(x_{J_1}, \cdots, x_{J_m}|r_i)=\binom{i}{m}=\binom{r_iN}{m},$$

where the summation is over all selections of m integers $J_1, J_2, \cdots, J_m$ from the from the integers $1, 2, \cdots, N$.

When this equation is divided by $\binom{N}{m}$, the right side becomes

$$\frac{\binom{r_i N}{m}}{\binom{N}{m}}=\frac{\dfrac{(r_i N)!}{m!(r_i N-m)!}}{\dfrac{N!}{m!(N-m)!}},$$

and after cancellations the right member of the equation in Theorem III is obtained.

Proof of Theorem VI

The hypothesis implies that the $N\times N$ matrix, of which the kth row vector is $(\pi_1^{(k)}, \pi_2^{(k)}, \cdots, \pi_K^{(k)})$, is non-singular. It follows that the system of equations $\pi_1^{(k)}c_1+\pi_2^{(k)}c_2+\cdots\pi_N^{(k)}c_N=0$, $k=1, 2, \cdots, N$, in the "unknowns" $c_1, c_2, \cdots, c_N$, has the unique solutions $c_1=c_2=\cdots=c_N=0$. Let J and K be any two integers, $1\leqq J<K\leqq N$. The hypothesis states that $p(x_j=1)=p(x_K=1)$. Therefore by Eq. (3),

$$\sum_{i=1}^{N} p(x_J=1\,|\,r_i)\pi_i^{(k)}=\sum_{1}^{N} p(x_K=1\,|\,r_i)\pi_i^{(k)},$$

or

$$\sum_{i=1}^{N} [p(x_J=1\,|\,r_i)-p(x_K=1\,|\,r_i)]\pi_i^{(k)}=0,$$

for $k=1, 2, \cdots, N$. Thus,

$$p(x_J=1\,|\,r_i)-p(x_K=1\,|\,r_i)=0,\ i=1, 2, \cdots, N.$$

The proof of the theorem is complete.

Suppose that in a set S of N specifications of the probability vector $\boldsymbol{\pi}$ of the proportion r of areal coverage, the kth vector $\boldsymbol{\pi}^k$ gives preferential probability to the value $r=r_k=k/N$ in the sense defined in the text. Then the matrix of which the row vectors are the vectors S with first components deleted has a dominant diagonal. Therefore, by a well-known theorem [for example, see Varga (1962, p. 23)] this matrix is non-singular, so its row vectors are linearly independent and the hypothesis of Theorem VI is satisfied.

Acknowledgments. The author is indebted to a number of professional meteologists who read a first draft of this paper and whose detailed comments greatly improved the relevance of the paper to the problems of weather forecasting. Especial thanks are due to R. H. Simpson, C. F. Roberts, and L. A. Hughes of the U. S. Weather Bureau. Prof. Frederick Sanders of the Massachusetts Institute of Technology and Prof. R. L. Winkler of Indiana University also made helpful comments. Conversations with Joseph Pelissier of the National Hurricane Center of Miami were valuable. A careful referee report by E. S. Epstein was much appreciated.

REFERENCES

Brier, G. W., 1950: Verification of forecasts expressed in terms of probability. *Mon. Wea. Rev.*, **78**, 1–3.

Cook, W. E., 1906: Weighting forecasts. *Mon. Wea. Rev.*, **34**, 274–275.

Dickey, W. W., 1965: Probability forecasting. U. S. Department of Commerce, ESSA, Weather Bureau Southern Region, Tech. Memo. No. 6, 14 pp.

Epstein, E. S., 1966a: Point and area precipitation probabilities. *Mon. Wea. Rev.*, **94**, 595–598.

——, 1966b: Quality control for probability forecasts. *Mon. Wea. Rev.*, **94**, 487–493.

Feller, William, 1957: *An Introduction to Probability Theory and Its Applications*. Second Edition, New York, John Wiley and Sons, Inc., 461 pp.

Hallenbeck, Cleve, 1920: Forecasting precipitation in percentages of probability. *Mon. Wea. Rev.*, **48**, 645–647.

Hughes, L. A., 1965: On the probability forecasting of the occurrence of precipitation. U. S. Department of Commerce, ESSA Tech. Note 20-CR-3, 36 pp.

Kolmogorov, A. N., 1950: *Foundations of the Theory of Probability*. Translated by Nathan Morrison, New York, Chelsea Publishing Co., 70 pp.

Parzen, Emanuel, 1962: *Stochastic Processes*. San Francisco, Holden-Day, Inc., 464 pp.

Root, H. E., 1962: Probability statements in weather forecasting. *J. Appl. Meteor.*, **1**, 163–168.

Sanders, Frederick, 1963: On subjective probability forecasting. *J. Appl. Meteor.*, **2**, 191–201.

Varga, R. S., 1962: *Matrix Iterative Analysis*. Englewood Cliffs, N. J., Prentice-Hall, Inc., 322 pp.